Molecular evolution: **Aminoacyl-tRNA synthetases on the loose**

Alan M. Weiner

Modified versions – paralogs – of the catalytic domain of at least three different aminoacyl-tRNA synthetases have been found to serve catalytic or regulatory roles in other reactions. These findings suggest that the first modern tRNA-synthetases could have been derived from amino-acid biosynthetic enzymes.

Address: Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, Connecticut 06520-8024, USA.
E-mail: weiner@biomed.med.yale.edu

Current Biology 1999, 9:R842–R844

0960-9822/99/$ – see front matter

It would have been much simpler if every protein, once invented, were held sacred and used exclusively for the intended purpose. Instead, evolution shamelessly cannibalizes existing polypeptides for new purposes, mixing and matching whole proteins and protein domains with reckless abandon. Sometimes old proteins are recycled intact. Aconitase, a citric acid cycle enzyme, has also been put to work as an iron-responsive RNA-binding protein that controls translation and stability of ferritin and transferrin receptor mRNAs [1]. Other enzymes of intermediary metabolism, such as aldehyde dehydrogenase and transketolase, double as crystallins in the vertebrate eye [2]. Especially shocking was the discovery that human tyrosyl-tRNA synthetase is secreted during apoptosis, then cleaved by extracellular proteases into two fragments with cytokine activity [3].

Other times, old proteins require tinkering before reuse. This is the case for the twenty aminoacyl-tRNA synthetases which couple — or 'charge' — amino acids onto the cognate tRNAs. Surprisingly, the twenty synthetases can be divided into two classes of ten each, the class I proteins being based on the Rossmann dinucleotide-binding fold, and the class II proteins on an unusual antiparallel β-sheet structure [4,5]. As all aminoacyl-tRNA synthetases use a similar, two-step charging mechanism — formation of an aminoacyl-adenylate followed by transfer of the activated amino acid to the 2′ or 3′ end of tRNA — it seems certain that ancestral class I and class II enzymes each gave rise to ten offspring.

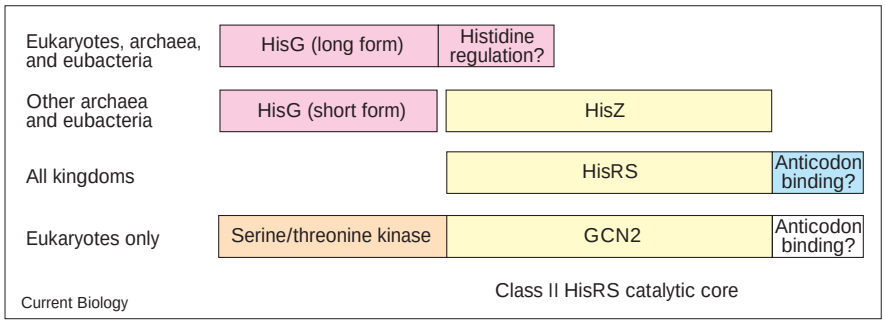
What cannot yet be explained is the existence of two unrelated classes of aminoacyl-tRNA synthetases. Once nature had gone to the trouble of inventing the first aminoacyl-tRNA synthetase, one might have thought that all subsequent synthetases would be derived from it. In fact, the most plausible explanation for two classes of aminoacyl-tRNA synthetases seems somewhat farfetched: the last common ancestor of all living organisms might have been a chimera of two ancient life forms, one with class I and the other with class II enzymes. This could account for what might be a functionally complete complement of amino acids in each of the two classes — glutamate, cysteine, tyrosine and arginine in class I, and aspartate, serine, phenylalanine and histidine in class II — with only lysyl-tRNA synthetase represented in both classes [6].

Just as the original class I and class II enzymes multiplied and diversified, giving rise to two families of aminoacyl-tRNA synthetases, so the catalytic domains appear to have broken loose to perform unexpected catalytic and regulatory functions. Four such instances are known. Asparagine synthetase (AsnA), which converts aspartate to asparagine by means of a transient aspartyl-adenylate (aspartyl-AMP) intermediate, is homologous to the catalytic domains of the aspartyl-tRNA and asparaginyl-tRNA synthetases. AsnA thus exploits the first step of the normal aspartatyl/asparaginyl-tRNA charging reaction to activate aspartate for transamidation. Similarly, the biotin repressor BirA, a class II homolog possibly related to serinyl-tRNA synthetase, synthesizes biotinyl-adenylate (bio-AMP), an activated intermediate for biotin transfer which may also be a positive allosteric effector of BirA DNA binding.

The eukaryotic kinase GCN2, which phosphorylates the translation initiation factor eIF2α, consists of a serine/threonine kinase module fused to a carboxy-terminal domain homologous to the catalytic domain of histidinyl-tRNA synthetase (Figure 1). The kinase activity of GCN2 is regulated by binding of uncharged tRNA to the histidinyl-tRNA synthetase domain. AsnA, BirA and GCN2 thus all deploy an aminoacyl tRNA synthetase catalytic domain to bind one or more amino-acyl tRNA synthetase substrates (amino acid, ATP or tRNA). In AsnA and BirA, the aminoacyl tRNA synthetase catalytic domain performs a catalytic function; in GCN2, the function appears to be strictly regulatory. Notably, but as yet inexplicably, all these moonlighting aminoacyl tRNA synthetase catalytic domains — those of the aspartate, asparagine, serine and histidine charging enzymes — belong to class II.

Sissler *et al.* [7] have now reported the fourth, and perhaps most puzzling, instance of a freelance aminoacyl-tRNA synthetase catalytic domain. The first step of histidine

Figure 1



The long form of HisG found in archaea, eukaryotes and some eubacteria is about 300 amino acid residues long. The short form of HisG, found in other eubacteria, is only 200 residues long as it lacks the carboxy-terminal region of the long form. The short form of HisG is inactive alone, but fully active when complexed with the HisZ protein from the same organism. HisZ is homologous to the catalytic core of the class II histidyl-tRNA synthetase (HisRS) and may confer histidine regulation on HisG. Similarly, binding of uncharged $tRNA^{His}$ to a modified histidyl-tRNA synthetase catalytic domain regulates the serine/threonine kinase activity of the eukaryotic enzyme GCN2 (only the kinase and histidyl-tRNA synthetase-related regions are shown).

biosynthesis is catalyzed by HisG (Figure 2). Curiously, HisG comes in two flavors, a long form, of about 300 residues, found in archaea, eukaryotes and eubacteria, and a carboxy-terminally truncated short form, of about 200 residues, found in all other prokaryotic histidine prototrophs. A powerful clue to the significance of the long and short forms of HisG came from the fertile field of comparative genomics: all prokaryotes having the short form of HisG also contain another protein (of about 300 residues), known as HisZ. As judged by genetic complementation, the short form of HisG is inactive alone, but fully active when complexed with HisZ protein. Thus, the short form of HisG becomes functionally equivalent to the long form when it is complexed with HisZ, although HisZ exhibits no obvious homology to the carboxyl terminus of the long form of HisG.

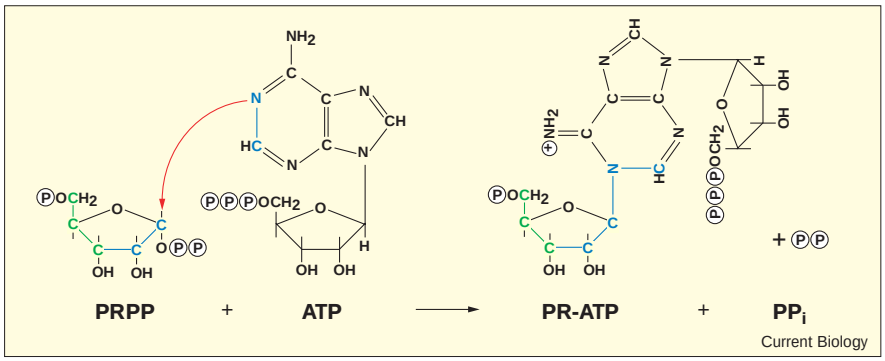
More remarkably, as first recognized by Renault, Delorme and colleagues [7], HisZ is homologous to the catalytic core of histidinyl-tRNA synthetase (Figure 1). Yet this catalytic core is not used catalytically for adenylation or amino-acylation: key residues involved in histidine binding are conserved between histidinyl-tRNA synthetase and HisZ, but those involved in catalysis are not, suggesting that binding of histidine to HisZ may allosterically regulate HisG. Consistent with this view, mutations in the histidine binding site of HisZ abolish histidine regulation of the short form of HisG (M. Sissler and C. Francklyn, personal communication).

So where did the long and short forms of HisG come from? According to an exhaustive phylogenetic analysis, HisZ is found in a divergent set of eubacteria and archaea, and appears to be monophyletic (J. P. Bond and C. Francklyn, personal communication). The simplest (though not the only) hypothesis to explain these data is that early prokaryotes contained HisZ and the short form of HisG; multiple, independent occurrences of the horizontal transfer from eukaryotes of the gene encoding the long form of HisG would then have driven HisZ (and the short form of HisG) out of numerous prokaryotic lineages. The long and short forms of HisG would therefore be paralogous not orthologous — that is, they were derived by an early gene duplication, not by speciation. After decades of eclipse, horizontal gene transfer has recently been coming into style. If confirmed, this could be the most widespread horizontal transfer event ever documented.

The big question, as in most evolutionary mysteries, is what came first. The simplest scenario is that the catalytic cores of three different preexisting class II synthetases were borrowed for new purposes — that of aspartyl/asparaginyl-tRNA synthetase by AsnA, that of histidinyl-tRNA synthetase by HisZ and GCN2, and that of a serinyl-tRNA synthetase-like enzyme by BirA. This scenario makes sense because aminoacyl-tRNA synthetases are required for translation, and a functional translation apparatus must have predated the origin and diversification of complex protein domains. The main trouble here is that two of the three borrowed catalytic cores participate in amino-acid biosynthesis — AsnA in building asparagine and HisZ in building histidine — and without amino acids there could be no protein synthesis. Is this another chicken-and-the-egg problem? The good news is we have lots of wiggle room; the bad news is we do not have answers.

In one scenario, translation got off to a fine start without asparagine or histidine, and these two amino acids were added later. This would finesse the problem because preexisting aminoacyl-tRNA synthetases could supply whatever class II catalytic domains were needed. In another scenario, asparagine and histidine were made by different pathways when translation was young; these ancient pathways were then replaced by modern pathways involving AsnA and HisZ, which arose after the

Figure 2



The biosynthesis of histidine begins with nucleophilic attack (red arrow) of the adenine ring of ATP on activated ribose (PRPP) to generate N′-5′-phosphoribosyl-ATP (PR-ATP). This reaction is catalyzed by HisG in archaea, eukaryotes and some eubacteria, but by a complex of HisG and HisZ in other eubacteria (also see Figure 1). Subsequent steps in this costly biosynthetic pathway open the six-membered adenine ring, and construct the imidazole side chain of histidine from atoms originally belonging both to ATP and PRPP (blue); the remaining histidine atoms (green) are derived from PRPP.

invention of aspartyl/asparaginyl-tRNA synthetase and histidine-tRNA synthetase, respectively.

In a more radical scenario favored by Sissler *et al.* [7], the class II catalytic core would have arisen to synthesize amino acids that were useful *before* the advent of a modern translation apparatus. Addition of an RNA-binding domain to this ancestral class II catalytic core would then allow transfer of the activated amino acid to the 2′ or 3′ hydroxyl of tRNA, either for primitive translation or for tRNA-dependent amino-acid biosynthesis, as still occurs today for asparagine, glutamine and selenocysteine [8]. Thus, it is not inconceivable that the catalytic core of class II aminoacyl-tRNA synthetases was invented as an enzyme of intermediary metabolism to build or activate amino acids; these would then be borrowed to serve a role in protein synthesis. Indeed this might help to explain the existence of two classes of tRNA synthetases. Perhaps Nature originally invented class I synthetases, and then discovered that class II synthetases could be fashioned from enzymes originally designed to build amino acids.

The most radical scenario invokes the ‘RNA World’. Very few would take issue with Crick [9] that all essential components of the original translation apparatus must have been made of RNA (or something like it) because complex proteins, like the class II catalytic core, could not have existed before the advent of translation. Thus the first aminoacyl-tRNA synthetases must have been made of RNA, a speculation now supported by many elegant proof-of-principle experiments [10]. Just as ancient RNAs persist today as essential catalytic components of the ribosome and spliceosome, so these ancient aminoacyl-tRNA synthetases made of RNA may have persisted as the RNA World gradually transformed itself into an RNP (ribonucleoprotein) World. Then neither aminoacyl-tRNA synthetases nor amino acid biosynthesis would have to come first; RNA could have done both jobs until proteins began to help out.

Acknowledgements

I thank Chris Francklyn and Marie Sissler for sharing thoughts and information in advance of publication.

References

1. Rouault T, Klausner R: Regulation of iron metabolism in eukaryotes. *Curr Top Cell Regul* 1997, 35:1-19.
2. Piatigorsky J: Multifunctional lens crystallins and corneal enzymes. More than meets the eye. *Ann N Y Acad Sci* 1998, 842:7-15.
3. Wakasugi K, Schimmel P: Two distinct cytokines released from a human aminoacyl-tRNA synthetase. *Science* 1999, 284:147-151.
4. Eriani G, Delarue M, Poch O, Gangloff J, Moras D: Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature* 1990, 347:203-206.
5. Delarue M, Moras D: The aminoacyl-tRNA synthetase family: modules at work. *Bioessays* 1993, 15:675-687.
6. Ibba M, Losey HC, Kawarabayasi Y, Kikuchi H, Bunjun S, Söll D: Substrate recognition by class I lysyl-tRNA synthetases: a molecular basis for gene displacement. *Proc Natl Acad Sci USA* 1999, 96:418-423.
7. Sissler M, Delorme C, Bond J, Ehrlich SD, Renault P, Francklyn C: An aminoacyl-tRNA synthetase paralog with a catalytic role in histidine biosynthesis. *Proc Natl Acad Sci USA* 1999, 96:8985-8990.
8. Curnow AW, Tumbula DL, Pelaschier JT, Min B, Söll D: Glutamyl-tRNA^{Gln} amidotransferase in ***Deinococcus radiodurans*** may be confined to asparagine biosynthesis. *Proc Natl Acad Sci USA* 1998, 95:12838-12843.
9. Crick FH: The origin of the genetic code. *J Mol Biol* 1968, 38:367-379.
10. Yarus M, Illangasekare M: Aminoacyl-tRNA synthetases and self-acylating ribozymes. In *The RNA World II*. Edited by Gesteland RF, Cech TR, Atkins JF. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 1999:183-196.